

Release 3.1A John F. Collins, Bioinformatics Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MP-SEARCH protein - protein database search, using Smith-Waterman algorithm

Thu Jan 27 17:33:59 2000; Maspar time 15.52 Seconds

496.003 Million cell updates/sec

Tabular output not generated.

Title: >US-09-397-538-1
Description: (1-141) from US09397538.pep
Perfect Score: 1097
Sequence: 1 MVHVAVSLCLPMBRSERLF.....ALVPSIVLDLGLCRYPD 141

Scoring table: PAM 150
Gap 11

Searched: 179066 seqs, 54579741 residues

Post-processing: Minimum Match 08
Listing first 45 summaries

Database: spiremb9
1:sp-archaea 2:sp-bacteria 3:sp-fungi 4:sp-human
5:sp-invertebrate 6:sp-mammal 7:sp-mhc 8:sp-organelle
9:sp-phage 10:sp-plant 11:sp-rodent 12:sp-unclassified
13:sp-vertebrate 14:sp-virus

Statistics: Mean 44.978; Variance 132.607; scale 0.339

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	Pred. No.
1	125	11.4	539	5	F40F9.5 PROTEIN.	9.36e-02
2	121	11.0	552	2	PUTATIVE PHOSPHATE PER	2.32e-01
3	120	10.9	380	8	CYTOCHROME B.	2.91e-01
4	116	10.6	488	8	NADH DEHYDROGENASE SUB	7.08e-01
5	116	10.6	1055	3	P-TYPE ATPASE 1.	7.08e-01
6	115	10.5	472	2	HOMOLOGUE OF MULTIDRUG	8.83e-01
7	112	10.2	256	14	HYPOTHETICAL 29.9 KD P	1.70e+00
8	112	10.2	386	2	HYPOTHETICAL 43.8 KD P	1.70e+00
9	110	10.0	305	8	YCF5 PROTEIN.	2.61e+00
10	110	10.0	319	5	F09P3.1 PROTEIN.	2.61e+00
11	109	9.9	381	8	CYTOCHROME B.	3.24e+00
12	109	9.9	693	8	NADH DEHYDROGENASE (FR	3.24e+00
13	107	9.8	125	2	HYPOTHETICAL 13.1 KD P	4.95e+00
14	108	9.8	192	8	CYTOCHROME B (FRAGMENT	4.01e+00
15	108	9.8	300	8	CYTOCHROME C (FRAGMENT	4.01e+00
16	108	9.8	311	8	CYTOCHROME C (FRAGMENT	4.01e+00
17	107	9.8	311	8	CYTOCHROME C (FRAGMENT	4.95e+00
18	108	9.8	379	8	CYTOCHROME B.	4.01e+00
19	108	9.8	381	8	CYTOCHROME B.	4.01e+00
20	108	9.8	381	8	CYTOCHROME B LIGHT STR	4.01e+00

21	107	9.8	381	8	P92717	CYTOCHROME B.	4.95e+00
22	107	9.8	381	8	C35545	CYTOCHROME B.	4.95e+00
23	107	9.8	414	2	007867	PUTATIVE POLYSACCHARID	4.01e+00
24	108	9.8	475	2	066589	NITRATE TRANSPORTER	4.01e+00
25	108	9.8	675	8	032213	NADH DEHYDROGENASE SUB	4.01e+00
26	107	9.8	685	8	019822	NADH DEHYDROGENASE (FR	4.95e+00
27	107	9.8	696	8	012058	NADH DEHYDROGENASE SUB	4.95e+00
28	108	9.8	702	8	019823	NADH DEHYDROGENASE (FR	4.01e+00
29	106	9.7	192	8	035642	CYTOCHROME B (FRAGMENT	6.11e+00
30	106	9.7	368	1	058546	368AA LONG HYPOTHETICA	6.11e+00
31	106	9.7	379	8	P92657	CYTOCHROME B.	6.11e+00
32	105	9.7	380	8	047545	APC/CYTOCHROME B.	6.11e+00
33	105	9.7	684	8	019831	NADH DEHYDROGENASE (FR	6.11e+00
34	106	9.7	685	8	019825	NADH DEHYDROGENASE (FR	6.11e+00
35	106	9.7	699	8	019854	NADH DEHYDROGENASE-LIK	6.11e+00
36	106	9.7	699	8	019856	NADH DEHYDROGENASE-LIK	6.11e+00
37	106	9.7	699	8	019853	NADH DEHYDROGENASE-LIK	6.11e+00
38	106	9.7	699	8	019855	NADH DEHYDROGENASE-LIK	6.11e+00
39	106	9.7	702	8	019830	NADH DEHYDROGENASE (FR	6.11e+00
40	106	9.7	710	8	019824	NADH DEHYDROGENASE (FR	6.11e+00
41	105	9.6	197	1	028693	HYPOTHETICAL 21.8 KD P	7.54e+00
42	105	9.6	313	5	045426	F31E9.2 PROTEIN.	7.54e+00
43	105	9.6	348	8	095935	CYTOCHROME B (FRAGMENT	7.54e+00
44	105	9.6	382	8	035047	CYTOCHROME B LIGHT STR	7.54e+00
45	105	9.6	382	8	035518	CYTOCHROME B LIGHT STR	7.54e+00

ALIGNMENTS

RESULT 1
ID G20336 PRELIMINARY; PRT; 539 AA.
AC 020336;
CT (1-NOV-1996 (TREMBLREL. 01, CREATED)
DT 11-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
FF 01-JAN-1999 (TREMBLREL. 09, LAST ANNOTATION UPDATE)
DE F40F9.5 PROTEIN.
GG F40F9.5.
CC CAEORHABDITIS ELEGANS.
GG ENKAROTA; METAZOA; NEMATODA; SECERNENTEA; RHABDITIDA; RHABDITIDA;
CC RHABDITIDA; RHABDITIDEA; RHABDITIDEA; PELODERINHA; CAENORHABDITIS.
AN (1)
NP SEQUENCE FROM N.A.
RA BERKS M.;
RT SUBMITTED (APR-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
RP [2]
K2 SEQUENCE FROM N.A.
RX MEDLINE; 94150718.
YA WILSON R., AINSOUGH R., ANDERSON K., BAYNES C., BERKS M.,
YA BOUTFIELD J., BURTON J., CONNELL M., COPESEY T., COOPER J., COULSON A.,
YA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,
YA SANDER A., GREEN P., HAKINS T., HILLIER L., JIER M., JOHNSTON L.,
YA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LARREILLE P.,
YA LIGHTNING J., LLOYD C., MCMURRAY E., MORTIMORE B., O'CALLAGHAN M.,
YA PARSONS J., PERCY C., RIKKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,
YA SALLON N., SMITH A., SONNHAMMER E., STADEN R., SULTON J.,
YA THIERRY-MIEL J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSOHN R.,
YA WATSON A., WEINSTOCK L., WILKINSON-SPRAT J., WOHLDMAN P.,
RT 3.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT flegans.
FU NATURE 368:32-38(1994).
DR EMBL; Z70753; E1346477;
SQ SEQUENCE 539 AA; 60947 MW; BDAE1002 CRC32;

Query Match 11.48; Score 125; DB 5; Length 539;
Best Local Similarity 38.98; Pred. No. 9.36e-02;
Matches 21; Conservative 18; Mismatches 12; Indels 3; Gaps 3;

DB 193 FVCSICLNGTILYSLAVALPAITSVERHALSYVCATIORALISITIVAFV 245
QY 47 YI-SFVMSLGL-SLAVALSIRSV-NALNWRVRSFIOSTIGVALLISTFHV 97
RESULT 2

ID 050684 PRELIMINARY; PRT; 552 AA.
 AC 050684;
 DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
 DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
 DE 01-FEB-1997 (TREMBLREL. 02, LAST ANNOTATION UPDATE)
 DE PUTATIVE PHOSPHATE PERMEASE CY339.29C.
 GN MTCY339.29C.
 OS MYCOBACTERIUM TUBERCULOSIS.
 OC BACTERIA; FIRMICUTES; ACTINOBACTERIA; ACTINOBACTERIDAE;
 OC ACTINOMYCETALES; CORYNEBACTERIINAE; MYCOBACTERIACEAE; MYCOBACTERIUM.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-H37RV;
 RA BADCOCK K., CHURCHER C.M., BARRELL B.G., RAJANDREAN M.A., WALSH S.V.;
 RL SUBMITTED (JUL-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -1- SIMILARITY: BELONGS TO THE PHO-4 FAMILY OF TRANSPORTERS.
 DR EMBL: 277163; E255063.
 KM HYPOTHETICAL PROTEIN; TRANSPORT; TRANSMEMBRANE; INNER MEMBRANE.
 FT TRANSMEM 38 POTENTIAL.
 FT TRANSMEM 69 POTENTIAL.
 FT TRANSMEM 107 127 POTENTIAL.
 FT TRANSMEM 146 166 POTENTIAL.
 FT TRANSMEM 178 198 POTENTIAL.
 FT TRANSMEM 213 233 POTENTIAL.
 FT TRANSMEM 326 346 POTENTIAL.
 FT TRANSMEM 360 380 POTENTIAL.
 FT TRANSMEM 389 409 POTENTIAL.
 FT TRANSMEM 437 457 POTENTIAL.
 FT TRANSMEM 472 492 POTENTIAL.
 FT TRANSMEM 493 513 POTENTIAL.
 FT TRANSMEM 526 546 POTENTIAL.
 SQ SEQUENCE 552 AA; 58789 MW; 3164CB08 CRC32;
 Query Match 11.08; Score 121; DB 2; Length 552;
 Best Local Similarity 28.68; Pred. No. 2,32e-01;
 Matches 26; Conservative 21; Mismatches 39; Indels 5; Gaps 4;
 Db 342 LKGFQKMLGLTTNNYFIAMGAAY-WMAFFIAKTLRG-ESLSRSTF--LMSMQ 397
 QY 46 MYISFGIMSLGLSLAVTSIPSVSNALNMRE-FSFIQSLGYALLSTFHVLYGKMR 104
 Db 398 VFTASGFASHGSDINAIQFPAILDVLR 428
 QY 105 AFEERYRYTPPNFVLALVPSIVLIDLQ 135
 F T 3
 I O63585 PRELIMINARY; PRT; 380 AA.
 AC 063585;
 DT 01-AUG-1998 (TREMBLREL. 07, CREATED)
 DT 01-AUG-1998 (TREMBLREL. 07, LAST SEQUENCE UPDATE)
 DT 01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
 DE CYTOCHROME B.
 OS FLOROMETRA SERRATISSIMA.
 OC MITOCHONDRION.
 OC EUKARYOTA; METAZOA; ECHINODERMATA; CRINOZOA; CRINOIDEA; ARTICULATA;
 OC COMATULIDA; MACROPHREATA; ANTEDONIDAE; FLOROMETRA.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA SCOURAS A., SMITH M.J.;
 RL SUBMITTED (FEB-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
 DE -1- CATALYTIC ACTIVITY: QH(2) + 2 FERROCYTOCHROME C = Q + 2
 CC -1- FERROCYTOCHROME C.
 CC COFACTOR: TWO HEME GROUPS
 CC (B562 AND B566) WHICH ARE NOT COVALENTLY BOUND TO THE PROTEIN
 CC (BY SIMILARITY).
 DR EMBL: AF049132; G2970421.
 DR PROSTATE; PS00192; CYTOCHROME_B_HEME.1.
 KM MITOCHONDRION; ELECTRON TRANSPORT; RESPIRATORY CHAIN; TRANSMEMBRANE;
 HEME.
 SQ SEQUENCE 380 AA; 43086 MW; 69C0C55A CRC32;

Query Match 10.98; Score 120; DB 8; Length 380;
 Best Local Similarity 26.08; Pred. No. 2,31e-01;
 Matches 25; Conservative 31; Mismatches 32; Indels 8; Gaps 7;
 Db 108 YVNEFMNNGVLLFLVMAAFYGVLPWGOMSEFGATV-ITVLSAIPPLGYELVQWV 166
 QY 47 YISFGIMSLGLSLAVTSIPSVSNALNMREFSFIQSLGYALLSTFHVLYGKMR 104
 Db 167 GGFSDNATLVRETF-HELLPEITIALSVYHLL 201
 QY 105 A-F--EEERY-RTTPPNFVLALVPSIVLIDLQ 136
 RESULT 4
 ID 021048 PRELIMINARY; PRT; 488 AA.
 AC 021048;
 DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
 DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
 DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
 DE NADH DEHYDROGENASE SUBUNIT 2 (EC 1.6.5.3) (NADH DEHYDROGENASE
 DE (UBIQUINONE)) (UBIQUINONE REDUCTASE) (TYPE I DEHYDROGENASE).
 DE (COMPLEX I DEHYDROGENASE).
 GN NAD2.
 OS DICTYOSTELIUM DISCOIDEUM (SLIME MOLD).
 OC MITOCHONDRION.
 OC EUKARYOTA; DICTYOSTELIIDA; DICTYOSTELIUM.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-AX3;
 RX MEDLINE: 97153471.
 RA OGAMA S., MATSUO K., ANGATA K., YANAGISAWA K., TANAKA Y.;
 RT "Group I" introns in the cytochrome c oxidase genes of Dictyostelium
 R1 discoidium: two related ORFs in one loop of a group-I intron, a
 RT cox1/2 hybrid gene and an unusually large cox3 gene.";
 CC CURR. GENET. 31:80-88(1997).
 CC -1- CATALYTIC ACTIVITY: NADH + UBIQUINONE = NAD(+) + UBIQUINOL.
 CC -1- COFACTOR: FAD; IRON-SULFUR.
 DR EMBL: D16579; D1021965;
 DR PFAM: PF00361; oxidored_q1; 1.
 KM OXIDOREDUCTASE; MITOCHONDRION.
 SC SEQUENCE 488 AA; 54841 MW; 4A899DB7 CRC32;
 Query Match 10.68; Score 116; DB 8; Length 488;
 Best Local Similarity 27.18; Pred. No. 7,08e-01;
 Matches 23; Conservative 31; Mismatches 27; Indels 4; Gaps 4;
 Db 268 YLNEVINKGLVLTTLVGTGSL-ROO-KVIRF-INYSAIVNSALLIFVGNNT 324
 QY 47 YISFGIMSLGLSLAVTSIPSVSNALNMREFSFIQSLGYALLSTFHVLYGKMR 106
 Db 325 ELIYSIYLDINYIIGLAVLNIIT 349
 QY 107 EEEYRYTPPNFVLALVPSIVL 130
 RESULT 5
 ID 013397 PRELIMINARY; PRT; 1055 AA.
 AC 013397;
 DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
 DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
 DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
 DE P-TYPE ATPASE 1.
 GN ENA1.
 OS DEBARYOMYCES OCCIDENTALIS.
 OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCCHAROMYCETALES;
 OC SACCCHAROMYCETACEAE; DEBARYOMYCES.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA BANDLOS M.A., RODRIGUEZ-NAVARRO A.;
 RL SUBMITTED (OCT-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
 DR EMBL: AF030860; G2623236;
 DR PFAM: PF00122; E1-E2_ATPase; 2.

SQ SEQUENCE 1055 AA; 116651 MW; 5A9F92CE CRC32;
 Query Match 10.6%; Score 116; DB 3; Length 1055;
 Best Local Similarity 24.0%; Pred. No. 7.08e-01;
 Matches 23; Conservative 37; Mismatches 26; Indels 10; Gaps 10;

DB 934 LRSFFKRPPELEPNMVKQALIDMDQ-FLFWSMG-AIYVFPVYIPIVINKVELHA 991
 21 LMAVQGVHANIENSM-NE-E-EVWRIEMYSFGIMSLGLSLAVTSPVSALNWMRE 77
 78 -FSFIQSTLGYVALLISTFHVLIYGMKRAFEVEEYR 112

RESULT 6 PRELIMINARY; PRT; 472 AA.
 AC 94422;
 DT 01-MAY-1997 (TREMBLREL. 03, CREATED)
 DT 01-NOV-1998 (TREMBLREL. 08, LAST SEQUENCE UPDATE)
 DE HOMOLOGUE OF MULTIDRUG RESISTANCE PROTEIN B, EMR, OF E. COLI.
 GN YCNB.
 OC BACILLUS SUBTILIS.
 OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; BACILLACEAE;
 OC BACILLUS.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-168 TRPC2;
 RX MEDLINE: 97124189.
 RA YAMANE K., KUMANO M., KURITA K.;
 RT "The 25 degrees-36 degrees region of the Bacillus subtilis
 chromosome: determination of the sequence of a 146 kb segment and
 identification of 113 genes.";
 RL MICROBIOLOGI 142:3047-3056(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-168;
 RX MEDLINE: 98044033.
 RA KUNST F., OGASAWARA N., MOSZER I., ALBERTINI A.M., ALLONI G.,
 AZEVEDO V., BERTERO M.G., BESSIERES P., BOLOTIN A., BORCHERT S.,
 BORRIS L., BOURSIER L., BRANS A., BRAUN M., BRIGNELL S.C., BRON S.,
 BROUILLET S., BRUSCHI C.V., CALDWELL B., CAPUANO V., CARTER N.M.,
 CHOI S.K., CODANI J.J., CONNERTON I.F., CUMMINGS N.J., DANIEL R.A.,
 DENOT F., DEVINE K.M., DUSTERHOFF A., EHRLICH S.D., EMERSON P.T.,
 ENTIAN K.D., ERINGTON J., FABRET C., FERRARI E., FOULGER D.,
 GARTY C., FUJITA M., FUJITA Y., FUMA S., GALIZZI A., GALLEON N.,
 GILBERT S.Y., GLASER P., GOFFEAU A., GOLIGHTLY E.J., GRANDI G.,
 GUISEPPI G., GUY B.J., HAGA K., HATECH J., HARWOOD C.R., HENAUT A.,
 HILBERT H., HOLSAPPEL S., HOSONO S., HULLO M.F., ITAYA M., JONES L.,
 JORIS B., KARAWATA D., KASAHARA Y., KLAER-BLANCARD M., KLEIN C.,
 KOBAYASHI Y., KOETTER P., KONINGSSTEIN G., KROGH S., KUMANO M.,
 KURITA K., LAPIDUS A., LARDINOIS S., LAUBER J., LAZAREVIC V.,
 LEE S.M., LEVINE A., LIU H., MASUDA S., MAUEL C., MEDIGUE C.,
 MEDINA N., MELLADO R.P., MIZUNO M., MOESTL D., NAKAI S., NOBACK M.,
 NOONE D., O'REILLY M., OGAMA K., OGIMARA A., OUDGA B., PARK S.H.,
 PARRO V., POHL T.M., PORTELELLA B., RABOPORT G., REY M., REYNOLDS S.,
 PRESCAN E., PUJIC P., PURNELLE B., RAPOPORT G., REY M., REYNOLDS S.,
 RIEGER M., RIVOLTA C., ROCHA E., ROCHÉ B., ROSE M., SABAIE Y.,
 SATO T., SCHNLAN E., SCHLECH S., SCHROETER R., SCOFONE F.,
 SEKICUCHI J., SEKOWSKA A., SEROR S.J., SEROR P., SHIN B.S., SOLDO B.,
 SOROKIN A., TACCONE E., TAKAGI T., TAKAHASHI H., TAKEMARU K.,
 TACHECHI M., TAMAKOSHI A., TANAKA T., TERSTRA P., TOGONOT A.,
 TOSATO V., UCHIYAMA S., VANDENBOL M., VANNIER F., VASSAROTI A.,
 VIARI A., WAMBOIT R., WEDLER E., WEDLER H., WEITZENGGER T.,
 WINTERS P., WIPAT A., YAMAMOTO H., YAMANE K., YASUMOTO K., YATA K.,
 YOSHIDA K., YOSHIKAWA H.F., ZUMSTEIN E., YOSHIKAWA H., DANCHIN A.;
 RT "The complete genome sequence of the gram-positive bacterium Bacillus
 subtilis.";
 RL NATURE 390:249-256(1997).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-168;

RA KUNST F., OGASAWARA N., YOSHIKAWA H., DANCHIN A.;
 RL SUBMITTED (NOV-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
 RN EMBL: D50453; G1805454;
 RN EMBL: 295106; E1182351;
 SQ SEQUENCE 472 AA; 50797 MW; 2836006E CRC32;

Query Match 10.5%; Score 115; DB 2; Length 472;
 Best Local Similarity 27.1%; Pred. No. 8.83e-01;
 Matches 23; Conservative 37; Mismatches 34; Indels 5; Gaps 4;

DB 145 LVSFAPAIPTYSGVNAVEAFSWRSLEFIILPFAVIDLIASLI-MKNVTLTKTQI-DI 202
 21 LMAVQGVHANIENSMNEE-VRIEMXI--SGIMSLGLSLAVTSPVSALNWMRE 77
 78 -FSFIQSTLGYVALLISTFHVLIYGM 102

RESULT 7 PRELIMINARY; PRT; 256 AA.
 AC 04638;
 DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
 DT 01-NOV-1998 (TREMBLREL. 08, LAST SEQUENCE UPDATE)
 DE HYPOTHETICAL 29.9 KD PROTEIN.
 GN SAIMIRINE HERPESVIRUS 2.
 OC VIRUSES; DSDNA VIRUSES, NO RNA STAGE; HERPESVIRIDAE;
 OC GAMMAHERPESVIRINAE.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-C-488;
 RX MEDLINE: 98037620.
 RA KNAPE A., HILDER C., THURAU M., WITTMANN S., HOFMANN H.,
 RA FLECKENSTEIN B., FICKENSCHER H.;
 RT "The superantigen-homologous viral immediate-early gene iel4/vsag in
 R herpesvirus saimiri-transformed human T cells.";
 RL J. VIROL. 71:9124-9133(1997).
 DR EMBL: Y13163; E332528;
 RT HYPOTHETICAL PROTEIN.
 SQ SEQUENCE 256 AA; 29889 MW; 5720E5A2 CRC32;

Query Match 10.2%; Score 112; DB 14; Length 256;
 Best Local Similarity 25.7%; Pred. No. 1.70e+00;
 Matches 27; Conservative 31; Mismatches 40; Indels 7; Gaps 7;

DB 101 IDPS-SRGQMLHNSNDLQGMQSFISKLGIOHFRISPONITDGNLFGVLPY 159
 32 IENSMNEEYV-RLEMYISF-GIMSLGLSLAVTSPVS-NALNMRSEF-IQSTLGY 87
 88 VALLISTFHVLIYGMKRAFEVEEYRPTPNFVLAVALPSVILD 132

RESULT 8 PRELIMINARY; PRT; 386 AA.
 AC 066545;
 DT 01-AUG-1998 (TREMBLREL. 07, CREATED)
 DT 01-AUG-1998 (TREMBLREL. 07, LAST SEQUENCE UPDATE)
 DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
 DE HYPOTHETICAL 43.8 KD PROTEIN.
 GN A0135.
 OS AUREFEX AEOIDICUS.
 OC BACTERIA; AQUIFICALES; AQUIFICACEAE; AQUIFEX.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-VF5;
 RX MEDLINE: 98196666.
 RA DECKERT G., WARREN P.V., GAASTERLAND T., YOUNG W.G., LENOX A.L.,
 RA GRAHAM D.E., OVERBERG R., SNEAD M.A., KELLER M., AUJAY M., HUBER R.,
 RA FELDMAN R.A., SHORT J.M., OLSON G.J., SKELLEN R.V.;
 RT "The complete genome of the hyperthermophilic bacterium Aquifex

```

RT aeolicus."
RL NATURE 392:353-358(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC DECKERT G., WARREN P.V., GASTERLAND T., YOUNG M.G., LENOX A.L.,
RA GRAHAM D.E., OVERBECK R., SNEAD M.A., KELLER M., AJAYI M., HUBER R.,
RL FELDMAN R.A., SHORT J.M., OLSON G.J., SWANSON R.V.
DR SUBMITTED (JUL-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
EMBL: AE00676; G2982892; .
KW HYPOTHEICAL PROTEIN.
SQ SEQUENCE 386 AA; 43793 MW; A4F7E62 CRC32;

Query Match
Best Local Similarity 25.98; Score 112; DB 2; Length 386;
Matches 21; Conservative 27; Mismatches 25; Indels 8; Gaps 7;

DB 2 LW-FSLFI-LSIAGSGFALLVAIARTPLADLP-PKY-FYALVGHVDSALIVGLYAF 57
QY 41 VRIEMKIVISFGISLGLSL-L-AVTSIPSVSNALNMRREFSFGISGLGY--ALLISTFHY 97
DB 58 LIFLWHRIFEKKENFASFLPA 78
QY 98 LLYGMKRAFEF-EYRYFTPP 117

RESULT 9
ID 019916 PRELIMINARY; PRT; 205 AA.
AC 019916;
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT 01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
DE YCEF5 PROTEIN.
GN YCEF5.
OS CYANIDIUM CALDARIUM (GALDIERIA SULPHURARIA).
OC CHLOROPLAST.
OC EUKARYOTA; RHODOPHYTA; BANGIOPHYCEAE; PORPHYRIDALES; PORPHYRIDACEAE;
OC CYANIDIUM.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-RK1;
RA GLOECKNER G., ROSENTHAL A., VALENTIN K.;
RL SUBMITTED (SEP-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: AF02186; G2465744; .
KW CHLOROPLAST.
SQ SEQUENCE 205 AA; 23273 MW; FBE22A8F CRC32;

Query Match
Best Local Similarity 34.88; Score 110; DB 8; Length 205;
Matches 24; Conservative 13; Mismatches 29; Indels 3; Gaps 3;

DB 121 QISKRTFHVALLVLLD-SVIGTVLTSPLLEIGSYFIDRLADTLILI-TFMISIVA 178
QY 43 RIEMYSIFGMSLGLSLAVTSIPSVSNALNMR-EFSFGISGLGYALLISTFHYLVIG 101
DB 179 GTDALGQY 187
QY 102 WKRAFEERY 110

RESULT 10
ID P90817 PRELIMINARY; PRT; 319 AA.
AC P90817;
DT 01-MAY-1997 (TREMBLREL. 03, CREATED)
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT 01-JAN-1999 (TREMBLREL. 09, LAST ANNOTATION UPDATE)
DE F09F3.1 PROTEIN.
GN F09F3.1.
OS CAENORHABDITIS ELEGANS.
OC EUKARYOTA; METAZOA; NEMATODA; SECERNITEA; RHABDITIA; RHABDITIDA;
OC RHABDITINA; RHABDITOIDEA; RHABDITIDAE; PELODERINAE; CAENORHABDITIS.
RN [1]
RP SEQUENCE FROM N.A.

```

```

RA THOMAS K.;
RL SUBMITTED (OCT-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [2]
RP SEQUENCE FROM N.A.
RC MEDLINE; 94150718.
RX WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,
RA BONFIELD J., BURTON J., CONNELL M., CORSEY T., COOPER J., COULSON A.,
RP CRATON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,
RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,
RA JONES N., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,
RA LIGHTNING J., LLOYD C., MCHURRAY A., MORTIMORE B., O'CALLAGHAN M.,
RA PARSONS J., PERCY C., RIEKEN L., ROOPRA A., SANDERS D., SHOWNKEEN R.,
RA SMAILDON N., SMITH A., SONNHAMMER E., STADEN R., SUSTON J.,
RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,
RA WATSON A., WEINSTOCK L., WILKINSON-SPOAT J., WOHLMAN P.;
RT "2.2 Mb of contigunse nucleotide sequence from chromosome III of C.
RL elegans."
RT NATURE 368:32-38(1994).
EMBL: 281056; E1345334; .
SQ SEQUENCE 319 AA; 36986 MW; BD9AD661 CRC32;

Query Match
Best Local Similarity 24.78; Score 110; DB 5; Length 319;
Matches 24; Conservative 26; Mismatches 40; Indels 7; Gaps 7;

DB 197 ITNFENIAIV-ITLMSNSGVNAESRRKRKRW-NTQCVIYDLSLQFDTINAY-IM 253
QY 44 IEMVIFGIMSLGLSL-L-A-VTSIPSVSNALNMRREFSFGISGLGYALLISTFHYLVIG 101
DB 254 W-RFYDAWYRFLEATLSILVSLDGFVMTYHOC 289
QY 102 WKRAFEERYRF-YTPNFVALVLPSTYIYLLDQLC 137

RESULT 11
ID P87410 PRELIMINARY; PRT; 381 AA.
AC P87410;
DT 01-MAY-1997 (TREMBLREL. 03, CREATED)
DT 01-MAY-1997 (TREMBLREL. 03, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE CYTOCHROME B.
CN CYTB.
OS MYRMECOBIUS FASCIAIUS.
OS MITOCHONDRIUM.
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; METATHERIA;
OC DASYROMORPHIA; MYRMECOBIIDAE; MYRMECOBIUS.
RN [1]
RP SEQUENCE FROM N.A.
RA KRAJEWSKI C., PAINTER J., DRISKELL A.C., BUCKLEY L., WESTERMAN M.;
RI SCI. NEW GUINEA 19:157-166(1995).
RN [2]
RP SEQUENCE FROM N.A.
RA KRAJEWSKI C., WESTERMAN M.;
RL SUBMITTED (DEC-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -1- CATALYTIC ACTIVITY: QH(2) + 2 FERRICYTOCHROME C - Q + 2
CC -1- FERROCYTOCHROME C.
CC -1- COFACTOR: TWO HEME GROUPS
CC (B562 AND B566) WHICH ARE NOT COVALENTLY BOUND TO THE PROTEIN
CC (BY SIMILARITY).
DR EMBL: U82329; G1785600; .
DR PROSITE: PS00192; CYTOCHROME_B_HEME; 1.
DR PFAM: PF00032; cytochrome_b_c; 1.
DR PFAM: PF00033; cytochrome_b_n; 1.
KW MITOCHONDRIUM; ELECTRON TRANSPORT; RESPIRATORY CHAIN; TRANSMEMBRANE;
KW HEME.
SQ SEQUENCE 381 AA; 42795 MW; DA0B368 CRC32;

Query Match
Best Local Similarity 25.88; Score 109; DB 8; Length 381;
Matches 23; Conservative 28; Mismatches 30; Indels 8; Gaps 8;

DB 114 NGVILLTVMTAFVGLPVGMSFWGATV-ITLLSAIFYIGSLVEMWVGFSYDK 172

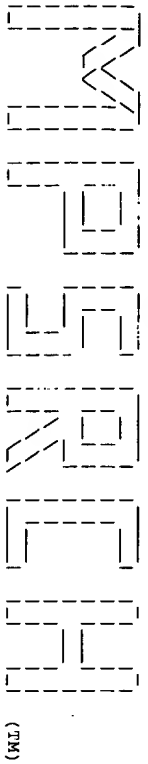
```

OY 54 SLGLSLAVTSIPSVSNALNMRFSIQTSLGVALLIST-FH-VLIYGMKRA-FE-E-108
 DB 173 ATLTREFAL-HEILPELITALYVHLFL 200
 OY 109 EYV-RFTYTPNFVALVLPISVILDLLOL 136
 RESULT 12
 ID 019828 PRELIMINARY; PRT; 595 AA.
 AC 019828;
 DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
 DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
 DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
 DE NADH DEHYDROGENASE (FRAGMENT).
 GN NDHF.
 OS EPISCIA SPHALERA.
 OG CHLOROPLAST.
 OC EUKARYOTA; VIRIDIPLANTAE; STREPTOPHYTA; EMBRIOPHYTA; TRACHEOPHYTA;
 OC EUPHYLOPHYTES; SPERMATOPHYTES; MAGNOLIOPHYTA; EUDICOTYLEDONS;
 OC ASTERIDAE; GENTIANANAE; LAMIALES; GESNERIACEAE; EPISCIA.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA SMITH J.F., KRESGE M.E., MOLLER M., CRONK Q.C.;
 RL SUBMITTED (JUL-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
 DR EMBL; AF013701; G2431894;
 DR PFAM; PF00361; oxidored_q1; 1.
 DR PFAM; PF00662; oxidored_q1_N; 1.
 DR PFAM; PF01010; oxidored_q1_C; 1.
 KW CHLOROPLAST.
 FT NON_TER 1
 FT NON_TER 1
 FT NON_TER 1
 SQ SEQUENCE 695 AA; 78608 MW; 5C8967A1 CRC32;
 Query Match 9.9%; Score 109; DB 8; Length 695;
 Best Local Similarity 28.4%; Pred. No. 3.24e+00;
 Matches 19; Conservative 22; Mismatches 22; Indels 4; Gaps 3;
 DB 76 EFGYLDPLTISMSMLITTVGIVNLYSDNYMAYDOGLRFFTYMSFFSTMLGLVTSN 135
 OY 77 EFSIQSTGLGV-ALLIST-FHVLIGWK-RAREEYRYRTPNFVALVLPISVILD 132
 DB 136 LIQIYIF 142
 OY 133 LLOLCRY 139
 RESULT 13
 ID 052329 PRELIMINARY; PRT; 125 AA.
 AC 052329;
 DT 01-JUN-1998 (TREMBLREL. 06, CREATED)
 DT 01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)
 DT 01-JUN-1998 (TREMBLREL. 06, LAST ANNOTATION UPDATE)
 DE HYPOTHETICAL 13.1 KD PROTEIN.
 OS SALMONELLA TYPHIMORIUM.
 OC BACTERIA; PROTEOBACTERIA; GAMMA SUBDIVISION; ENTEROBACTERIACEAE;
 OC SALMONELLA.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA GUNN J.S., LIM K.B., KRUEGER J., KIM K., GUO L., HACKETT M.,
 RA MILLER S.I.;
 RA MOL. MICROBIOL. 0:0-0(1997).
 DR EMBL; AF036677; G2921425;
 KW HYPOTHETICAL PROTEIN
 SQ SEQUENCE 125 AA; 13148 MW; 44B10BEE CRC32;
 Query Match 9.8%; Score 107; DB 2; Length 125;
 Best Local Similarity 27.2%; Pred. No. 4.95e+00;
 Matches 25; Conservative 27; Mismatches 35; Indels 5; Gaps 5;
 DB 8 ISVAIASLAQSLGFAMRLPSIAHPLAFISGLGAFNATIALFAGL-AGLYSVFCMQ 66
 OY 48 ISGIMSIGLST-LAVTSIPSVSNALNM-REF-SFIQSTGLGVALLISTFHVLIYGMK 104

DB 67 TLTLALNKAVALLSTSVLWVWASMLLPGLQ 98
 OY 105 AFEE-EYVRYTPNFVALVLPISVILDLLOL 135
 RESULT 14
 ID 037055 PRELIMINARY; PRT; 192 AA.
 AC 037055;
 DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
 DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
 DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
 DE CYTOCHROME B (FRAGMENT).
 OS PHOCENOIDEIS DALLI.
 GN MITOCHONDRION.
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; CETACEA;
 OC ODONTOCETI; PHOCOENIDAE; PHOCOENOIDEIS.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN-PD14;
 RA MEDLINE; 98360502.
 RA ROSEL F.E., HAYGOOD M.G., PERRIN W.F.;
 RA "phylogenetic relationships among the true porpoises
 (Cetacea: Phocoenidae)."
 FT MOL. PHYLOGENET. EVOL. 4:463-474(1995).
 CC -1- CATALYTIC ACTIVITY: OH(2) + 2 FERRICYTOCHROME C - Q + 2
 CC -1- FERRICYTOCHROME C.
 CC -1- COFACTOR: TWO HEME GROUPS
 CC (B562 AND B566) WHICH ARE NOT COVALENTLY BOUND TO THE PROTEIN
 (BY SIMILARITY).
 DR EMBL; U09679; G944853;
 DR EMBL; U09678; G944851;
 DR PROSITE; PS00192; CYTOCHROME_B_HEME; 1.
 DR PFAM; PF00033; cytochrome_b_N; 1.
 KW MITOCHONDRION; ELECTRON TRANSPORT; RESPIRATORY CHAIN; TRANSMEMBRANE;
 KW HEME.
 FT NON_TER 1
 FT NON_TER 1
 FT NON_TER 1
 SQ SEQUENCE 192 AA; 21618 MW; 95AFACFF CRC32;
 Query Match 9.8%; Score 108; DB 8; Length 192;
 Best Local Similarity 26.7%; Pred. No. 4.01e+00;
 Matches 24; Conservative 28; Mismatches 28; Indels 10; Gaps 9;
 DB 101 NIGVLLIMWATFVGVYVPMGMSFRGATV-ITNLSPAIYGNLY-ETWIGSPVD 158
 OY 54 SLGLSLAVTSIPSVSNALNMRFSIQTSLGVALLIST-FHVLIGWKRA-FE-E-108
 DB 159 KATLTREFAF-HEILPELITALYVHLFL 187
 OY 109 EYV-RFTYTPNFVALVLPISVILDLLOL 136
 RESULT 15
 ID 021828 PRELIMINARY; PRT; 300 AA.
 AC 021828;
 DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
 DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
 DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
 DE CYTOCHROME B (FRAGMENT).
 GN CYTB.
 OS CTENOSAURA HEMILOPHA.
 OG MITOCHONDRION.
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; LEPIDOSAURIA; SQUAMATA;
 OC IGUANIA; IGUANIDAE; IGUANINAE; CTENOSAURA.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA PERRIN F., CASE T.J.;
 RA EVOLUTION 51:206-219(1997).
 CC -1- CATALYTIC ACTIVITY: OH(2) + 2 FERRICYTOCHROME C - Q + 2
 CC -1- FERRICYTOCHROME C.
 CC -1- COFACTOR: TWO HEME GROUPS
 CC (B562 AND B566) WHICH ARE NOT COVALENTLY BOUND TO THE PROTEIN

(BY SIMILARITY).
CC EMBL: AF020254; G2444365; -
DR EMBL: AF020253; G2444363; -
DR PROSITE: PS00192; CYTOCHROME_B_HEME; 1.
DR PFAM: PF00032; cytochrome_b_C; 1.
DR PFAM: PF00033; cytochrome_b_N; 1.
KM MITOCHONDRION; ELECTRON TRANSPORT; RESPIRATORY CHAIN; TRANSMEMBRANE;
HEME.
FT NON_TER 1 1
FT NON_TER 300 300
SQ SEQUENCE 300 AA; 33590 MW; C6C69B4E CRC32;
Query Match 9.88; Score 108; DB 8; Length 300;
Best Local Similarity 29.28; Pred. No. 4.01e+00;
Matches 26; Conservative 24; Mismatches 31; Indels 8; Gaps 7;
Db 97 NLGVILLVMTAFVGYVLPWGMSPFGATV-ITNLLSATPYVGTTLVEMINGFSVDN 155
QY 54 SLGLSLAVTSIPSYSNALNMRFSFIOSTLGVALLIST-FH-VLIYGMKRA-F--EE 108
D 156 ATLTFEFTF-HLLLPVITIGAAVLAHLLEL 183
QY 109 EYX-REYTPNFFVLAIVLPSIVILDLQL 136

Search completed: Thu Jan 27 17:34:20 2000
Job time : 21 secs.



Release 3.1A John F. Collins, B.Occupational Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

Maspar - protein - protein database search, using Smith-Waterman algorithm
Run: Thu Jan 27 17:40:11 2000; Maspar time 7.52 Seconds
700.292 Million cell updates/sec
Tabular output not generated.

Title: >US-09-397-558-2
Description: (1-410) from US09397558.pep
Perfect score: 3201
Sequence: 1 MEPLPVVLAIKRSRYVLEAAV.....RDCRMGAGTGVAVGILV 410

Scoring table: PAM 150
Gap 11

Searched: 130303 seqs, 12848679 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: a-issued
1:5A_COMB 2:5B_COMB 3:PCT9_COMB 4:backfile1

Statistics: Mean 33.711; Variance 155.316; scale 0.217

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Re No.	Score	Query Match	Length	DB	ID	Description	Pred. No.
1	781	24.4	229	1	US-08-702-	Sequence 9, Applicatio	2.97e-58
2	123	3.8	334	1	US-08-559-	Sequence 2, Applicatio	3.33e-01
3	110	3.4	591	1	US-08-483-	Sequence 2, Applicatio	2.86e+00
4	110	3.4	591	1	US-08-484-	Sequence 2, Applicatio	2.86e+00
5	101	3.2	265	3	PCT-US95-0	Sequence 29, Applicati	1.21e+01
6	101	3.2	265	1	US-08-413-	Sequence 27, Applicati	1.21e+01
7	101	3.2	265	1	US-08-321-	Sequence 5, Applicatio	1.03e+01
8	102	3.2	362	2	US-08-902-	Sequence 4, Applicatio	1.03e+01
9	102	3.2	1618	1	US-07-853-	Sequence 2, Applicatio	1.03e+01
10	102	3.2	1805	1	US-08-087-	Sequence 8, Applicatio	1.94e+01
11	98	3.1	1452	2	US-08-449-	Sequence 4, Applicatio	3.09e+01
12	98	3.1	1452	2	US-08-449-	Sequence 4, Applicatio	3.09e+01
13	95	3.0	1452	2	US-08-652-	Sequence 4, Applicatio	3.09e+01
14	95	3.0	1452	2	US-08-652-	Sequence 4, Applicatio	3.09e+01
15	95	3.0	1452	2	US-08-652-	Sequence 4, Applicatio	3.09e+01
16	94	2.9	98	3	PCT-US93-0	Sequence 336, Applicat	3.60e+01
17	94	2.9	98	1	US-08-118-	Sequence 22, Applicati	4.89e+01
18	92	2.9	316	1	US-08-414-	Sequence 22, Applicati	4.89e+01
19	92	2.9	316	2	US-08-926-	Sequence 3, Applicatio	4.89e+01
20	92	2.9	317	2	US-08-790-	Sequence 3, Applicatio	4.89e+01
21	92	2.9	317	2	US-09-213-	Sequence 8, Applicatio	4.89e+01
22	92	2.9	317	2	US-08-619-	Sequence 8, Applicatio	4.89e+01
23	92	2.9	340	2	US-09-213-	Sequence 1, Applicatio	4.89e+01

24	92	2.9	340	2	US-08-790-	Sequence 1, Applicatio	4.89e+01
25	93	2.9	361	1	US-08-415-	Sequence 7, Applicatio	4.20e+01
26	94	2.9	593	1	US-08-296-	Sequence 2, Applicatio	3.60e+01
27	53	2.9	890	2	US-08-451-	Sequence 8, Applicatio	4.20e+01
28	93	2.9	890	1	US-08-472-	Sequence 8, Applicatio	4.20e+01
29	93	2.9	890	2	US-08-323-	Sequence 8, Applicatio	4.20e+01
30	90	2.8	96	3	PCT-US93-0	Sequence 340, Applicat	6.63e+01
31	50	2.8	96	1	US-08-118-	Sequence 340, Applicat	6.63e+01
32	39	2.8	98	3	PCT-US93-0	Sequence 337, Applicat	7.71e+01
33	51	2.8	342	2	US-08-483-	Sequence 2, Applicatio	6.63e+01
34	90	2.8	370	2	US-08-031-	Sequence 2, Applicatio	6.63e+01
35	90	2.8	384	1	US-08-232-	Sequence 4, Applicatio	6.63e+01
36	90	2.8	384	3	PCT-US93-0	Sequence 3, Applicatio	6.63e+01
37	90	2.8	384	2	US-08-555-	Sequence 15, Applicati	6.63e+01
38	91	2.8	470	2	US-08-377-	Sequence 1, Applicatio	5.70e+01
39	89	2.8	1107	3	PCT-US96-0	Sequence 2, Applicatio	7.71e+01
40	89	2.8	1107	1	US-08-366-	Sequence 2, Applicatio	7.71e+01
41	89	2.8	1229	1	US-08-176-	Sequence 4, Applicatio	7.71e+01
42	89	2.8	1229	1	US-08-779-	Sequence 4, Applicatio	7.71e+01
43	89	2.8	1436	2	US-08-652-	Sequence 2, Applicatio	7.71e+01
44	89	2.8	1436	2	US-08-991-	Sequence 2, Applicatio	7.71e+01
45	89	2.8	2183	1	US-08-348-	Sequence 7, Applicatio	7.71e+01

ALIGNMENTS

RESULT 1
US-08-702-344-9 STANDARD: PRT: 229 AA.
xxxxxx
Sequence 5: Application US/08702344
Sequence 9: Application US/08702344
Patent No. 5723315
GENERAL INFORMATION:
APPLICANT: Jacobs, Kenneth
APPLICANT: McCoy, John
APPLICANT: Lavallee, Edward
APPLICANT: Racie, Lisa
APPLICANT: Merberg, David
APPLICANT: Treacy, Maurice
APPLICANT: Spaulding, Vikki
TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genetics Institute, Inc.
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version v1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/702,344
FILING DATE:
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Brown, Scott A.
REGISTRATION NUMBER: 32,724
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 498-8224
TELEFAX: (617) 876-5851
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 229 amino acids

CC TYPE: amino acid
CC STRANDEDNESS:
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 229 AA; 25907 MW; 284835 CN;
Query Match 24.4%; Score 781; DB 1; Length 229;
Best Local Similarity 57.7%; Pred. No. 2.97e-58;
Matches 82; Conservative 38; Mismatches 22; Indels 0; Gaps 0;
Db 84 MELAIASVRFVFFVSTVYATVTFSTFYHACDQPGFAVLCILSYDTLOXCDFLGSGA 143
OY 1 MFLPPVLAIRSRVLEAAVTFMTFSTFYHACDQPGIYVCINDYDLOFCDFLGSIM 60
Db 144 AIIWTTILCMARKLYLKYVLLFLGLTVIAMSLOLDRGMNMLGPCLFAFVIMASMMAYR 203
OY 61 SVWVTIVIMARLPQVKKQVLLGLAMLSMALQDLDRHGLMNLGSPSLFALGILATATVTR 120
Db 204 CGHRQCPTSMQKNAFYLLPG 225
OY 21 SVRRHCYPTWRRLFYLCPG 142
RESULT 2
ID US-08-559-524A-2 STANDARD: PRT: 334 AA.
AC xxxxxx
XX Sequence 2, Application US/08559524A
DE Sequence 2, Application US/08559524A
XX Sequence 2, Application US/08559524A
CC Patent No. 5871963
CC GENERAL INFORMATION:
CC APPLICANT: Conley, Pamela B.
CC APPLICANT: Jantzen, Hans-Michael
CC TITLE OF INVENTION: NOVEL PURINERGIC RECEPTOR
CC NUMBER OF SEQUENCES: 14
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: MORGAN, LEWIS & BOCKIUS LLP
CC STREET: 1800 M Street, N.W.
CC CITY: Washington
CC STATE: D.C.
CC COUNTRY: USA
CC ZIP: 20036-5869
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: floppy disk
CC OPERATING SYSTEM: IBM PC compatible
CC SOFTWARE: Patentin Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/559,524A
CC FILING DATE: 15-NOV-1995
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Adler, Reid G.
CC REGISTRATION NUMBER: 30,988
CC REFERENCE/DOCKET NUMBER: 04481-5010-00-US
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 202-467-7000
CC TELEFAX: 202-467-7176
CC INFORMATION FOR SEQ ID NO: 2:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 334 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 334 AA; 38657 MW; 618719 CN;
Query Match 3.8%; Score 123; DB 2; Length 334;
Best Local Similarity 30.3%; Pred. No. 3.33e-01;
Matches 23; Conservative 23; Mismatches 25; Indels 5; Gaps 5;

DE 93 VLCTSNRYVLNANLYTSLFL-TFT-STDRYLIIKYPREH-LLOKKEFALLISLAIVL 149
OY 7 VLAIIRSYVLEAAVTFMTFSTFYHACDQPGIYVCINDYDLOFCDFLGSIM 65
Db 150 VTLELLPILPLINPYI 165
OY 66 V-IMARLPQVKKQVLL 80
RESULT 3
ID US-08-483-094-2 STANDARD: PRT: 591 AA.
AC xxxxxx
XX Sequence 2, Application US/08483094
DE Sequence 2, Application US/08483094
XX Sequence 2, Application US/08483094
CC Patent No. 5763216
CC GENERAL INFORMATION:
CC APPLICANT: MOSCOW, Jeffrey A.
CC APPLICANT: COMAN, Kenneth H.
CC APPLICANT: DIXON, Kathy
CC APPLICANT: HE, Rui
CC TITLE OF INVENTION: A GENE ENCODING A HUMAN REDUCED FOLATE
CC CARRIER (RFC) AND METHODS FOR THE TREATMENT OF
CC TITLE OF INVENTION: METHOTREXATE-RESISTANT, TRANSPORT-DEFICIENT CANCER CELL
CC NUMBER OF SEQUENCES: 6
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Foley & Lardner
CC STREET: 3000 K Street, N.W., Suite 500
CC CITY: Washington
CC STATE: D.C.
CC COUNTRY: USA
CC ZIP: 20007-5109
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: floppy disk
CC OPERATING SYSTEM: IBM PC compatible
CC SOFTWARE: Patentin Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/483,094
CC FILING DATE: 07-JUN-1995
CC CLASSIFICATION: 514
CC ATTORNEY/AGENT INFORMATION:
CC NAME: BENT, Stephen A.
CC REGISTRATION NUMBER: 29,768
CC REFERENCE/DOCKET NUMBER: 40399/323/NIHD
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (202)672-5300
CC TELEFAX: (202)672-5399
CC INFORMATION FOR SEQ ID NO: 2:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 591 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 591 AA; 64921 MW; 1822494 CN;
Query Match 3.4%; Score 110; DB 1; Length 591;
Best Local Similarity 31.8%; Pred. No. 2.86e+00;
Matches 28; Conservative 22; Mismatches 29; Indels 9; Gaps 9;
Db 192 FITESSVVALFLKPKRSLFF-NRDDRRCETSSASELRNMPGGGKLGALRVACG-DS 249
OY 243 FLEFLGWSPILRR-RQAVFLFEDMDSFSTRTONSSRDELPF-PGHGE-LPEGLSPICIMES 299
Db 250 VLARMLRELG-DSLRR-R-POLRLMSLW 274
OY 300 FL-RIGAYAGTESLTKESLLOVMSLSW 326

RESULT 4
ID US-08-484-840-2 STANDARD: PRT: 591 AA.
AC xxxxxx
XX
XX
XX
XX
DE Sequence 2, Application US/08484840
CC Sequence 2, Application US/08484840
CC Patent No. 5716788
CC GENERAL INFORMATION:
CC APPLICANT: MOSCOW, Jeffrey A.
CC APPLICANT: COMAN, Kenneth H.
CC APPLICANT: DIXON, Kathy
CC APPLICANT: HE, Rul
CC TITLE OF INVENTION: A GENE ENCODING A HUMAN REDUCED FOLATE
CC TITLE OF INVENTION: CARRIER (RRC) AND METHODS FOR THE TREATMENT OF
CC TITLE OF INVENTION: METHOTREXATE-RESISTANT, TRANSPORT-DEFICIENT CANCER CELLS
CC NUMBER OF SEQUENCES: 6
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Foley & Lardner
CC STREET: 3000 K Street, N.W., Suite 500
CC CITY: Washington
CC STATE: D.C.
CC COUNTRY: USA
CC ZIP: 20007-5109
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/484,840
CC FILING DATE: 07-JUN-1995
CC CLASSIFICATION: 530
CC ATTORNEY/AGENT INFORMATION:
CC NAME: BENT, Stephen A.
CC REGISTRATION NUMBER: 29,768
CC REFERENCE/DOCKET NUMBER: 40399/324/NIHD
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (202)672-5300
CC TELEFAX: (202)672-5399
CC TELEX: 904136
CC INFORMATION FOR SEQ ID NO: 2:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 591 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 591 AA; 64921 MM; 1822494 CN;
Query Match 3.4%; Score 110; DB 1; Length 591;
Best Local Similarity 31.8%; Pred. No. 2.86e+00;
Matches 28; Conservative 22; Mismatches 29; Indels 9; Gaps 9;
Db 192 FLTSSVALFLKPKRSLFF-NRDDRRCETSSASELRNMPGGKGLHALRVACG-DS 249
QY 243 FLEFLIGWSPRLR-RQAVVFLEDMFSFRTONSSRDLEPF-PGHEE-LPECHESPCIMES 299
Db 250 VLARMLRELG-DSLRL-R-POLRLMSLW 274
QY 300 FL-RGAVAGTESLRTKESLLQVMSLSW 326
RESULT 5
ID PCT-US95-03776-29 STANDARD: PRT: 265 AA.
AC xxxxxx
XX
XX
XX
XX
XX
DL

XX
DE Sequence 29, Application PC/TUS9503776
XX
XX
XX
CC Sequence 29, Application PC/TUS9503776
CC GENERAL INFORMATION:
CC APPLICANT: AMGEN INC.
CC TITLE OF INVENTION: Compositions and Methods for Stimulating
CC TITLE OF INVENTION: Megakaryocyte Growth and Differentiation
CC NUMBER OF SEQUENCES: 34
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Amgen Inc.
CC STREET: 1840 Dehavilland Drive
CC CITY: Thousand Oaks
CC STATE: California
CC COUNTRY: USA
CC ZIP: 91320-1789
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US95/03776
CC FILING DATE:
CC CLASSIFICATION:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: COOK, Robert R.
CC REFERENCE/DOCKET NUMBER: A-290-C
CC INFORMATION FOR SEQ ID NO: 29:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 265 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 265 AA; 29189 MM; 380232 CN;
Query Match 3.2%; Score 101; DB 3; Length 265;
Best Local Similarity 33.3%; Pred. No. 1.21e+01;
Matches 19; Conservative 15; Mismatches 19; Indels 4; Gaps 4;
Db 170 SMSAPNLQVPG-PNPRPEDDTLTMNSVTLSTLTLTODPRSGHFLRNTRRLPA 225
QY 325 SWDAEPQSDQMSFPGRSVPVSTASFOR-WSLSWG-NOISRFQOR-LSNGLRLPS 378
RESULT 6
ID US-08-413-803-27 STANDARD: PRT: 265 AA.
AC xxxxxx
XX
XX
XX
XX
XX
DE Sequence 27, Application US/08413803
CC Sequence 27, Application US/08413803
CC Patent No. 5765581
CC GENERAL INFORMATION:
CC APPLICANT: Bartley, Timothy D.
CC APPLICANT: Bogensberger, Jakob M.
CC APPLICANT: Boseselmann, Robert A.
CC APPLICANT: Hunt, Pamela
CC APPLICANT: Kinstler, Olaf B.
CC APPLICANT: Samal, Babru B.
CC TITLE OF INVENTION: METHODS FOR TREATING MAMMALS WITH
CC TITLE OF INVENTION: MONO-PGGLYLATED PROTEINS THAT STIMULATE MEGAKARYOCYTE
CC NUMBER OF SEQUENCES: 34
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: AMGEN INC.
CC STREET: 1840 Dehavilland Drive
CC CITY: Thousand Oaks
CC STATE: California
CC COUNTRY: US

CC REGISTRATION NUMBER: 36,749
CC REFERENCE/DOCKET NUMBER: PF-0345 US
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 415-855-0555
CC TELEFAX: 415-845-4166
CC TELEX:
CC INFORMATION FOR SEQ ID NO: 5:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 362 amino acids
CC TYPE: amino acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC IMMEDIATE SOURCE:
CC LIBRARY: GenBank
CC CLONE: 1123105
SQ SEQUENCE 362 AA; 43020 MW; 726128 CN;

Query Match 3.2%; Score 102; DB 2; Length 362;
Best Local Similarity 23.1%; Pred. No. 1.03e+01;
Matches 9; Conservative 19; Mismatches 11; Indels 0; Gaps 0;

Db 151 SLIGSTFVDRSDFWOLMVHVTITFLSSMTINPVR 189
QY 85 AMLSMALQLDHRCGLMGLPSFLGLATAMTVRSYR 123

RESULT 9
ID US-07-853-913-4 STANDARD; PRT; 1618 AA.
XX
AC xxxxxx
DT
DE Sequence 4, Application US/07853913
CC Sequence 4, Application US/07853913
CC Patent No. 5338839
CC GENERAL INFORMATION:
CC APPLICANT: McKay, Ronald D.G.
CC APPLICANT: Lendahl, Urban
CC TITLE OF INVENTION: Nestin Expression As An Indicator of
CC TITLE OF INVENTION: Neuroepithelial Tumors
CC NUMBER OF SEQUENCES: 4
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
CC STREET: Two Millitia Drive
CC CITY: Lexington
CC STATE: Massachusetts
CC COUNTRY: U.S.A.
CC ZIP: 02173
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/07/853,913
CC FILING DATE: 19920319
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/660,412
CC FILING DATE: 22-FEB-1991
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/603,803
CC FILING DATE: 25-OCT-1990
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/201,762
CC FILING DATE: 02-JUN-1988
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/180,548
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Granahan, Patricia
CC REGISTRATION NUMBER: 32,227

CC REFERENCE/DOCKET NUMBER: MIT-4641AAAA
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 617-861-6240
CC TELEFAX: 617-861-9540
CC INFORMATION FOR SEQ ID NO: 4:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1618 amino acids
CC TYPE: AMINO ACID
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
SQ SEQUENCE 1618 AA; 176705 MW; 12580209 CN;

Query Match 3.2%; Score 102; DB 1; Length 1618;
Best Local Similarity 23.0%; Pred. No. 1.03e+01;
Matches 20; Conservative 23; Mismatches 43; Indels 1; Gaps 1;

Db 138 EALRVAAHEERY-GLNQAACAPRLPAPPPAPAPVEELARLGAMGAVRGIOERY 136
QY 310 ESLTKESLLQVWSLSMDAPSDQMSFPGRQSPVRSSTASFORWISLWGNQISRSQRL 369
Db 197 AHMETSLDQTRERLARAVOGAREVYLE 223
QY 370 SNGSLRLPQSORQRLGCAVLMRDCRMD 396

RESULT 10
ID US-07-853-913-2 STANDARD; PRT; 1805 AA.
XX
AC xxxxxx
DT
DE Sequence 2, Application US/07853913
CC Sequence 2, Application US/07853913
CC Patent No. 5338839
CC GENERAL INFORMATION:
CC APPLICANT: McKay, Ronald D.G.
CC APPLICANT: Lendahl, Urban
CC TITLE OF INVENTION: Nestin Expression As An Indicator of
CC TITLE OF INVENTION: Neuroepithelial Tumors
CC NUMBER OF SEQUENCES: 4
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
CC STREET: Two Millitia Drive
CC CITY: Lexington
CC STATE: Massachusetts
CC COUNTRY: U.S.A.
CC ZIP: 02173
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/07/853,913
CC FILING DATE: 19920319
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/660,412
CC FILING DATE: 22-FEB-1991
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/603,803
CC FILING DATE: 25-OCT-1990
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/201,762
CC FILING DATE: 02-JUN-1988
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/180,548
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Granahan, Patricia
CC REGISTRATION NUMBER: 32,227
CC REFERENCE/DOCKET NUMBER: MIT-4641AAAA

```

CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 617-861-6240
CC TELEFAX: 617-861-9540
CC INFORMATION FOR SEQ ID NO: 2:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1805 amino acids
CC TYPE: AMINO ACID
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 1805 AA; 198744 MW; 15616656 CN;
SO

Query Match 3.2%; Score 102; DB 1; Length 1805;
Best Local Similarity 26.7%; Pred. No. 1.03e+01;
Matches 16; Conservative 14; Mismatches 29; Indels 1; Gaps 1

Db 166 PHRTGPAPEVEDLARRIGEYWRGAVRDIQEVAMESSLGOARRLSQAVRGARECHLE 225
0y 338 PGR-OSPVKSTASFORRRSLSWGNOISRFSSQSLNSGRLPSSQRLRCGLAVLRDRCDMD 396

R 11
XX US-08-087-244A-8 STANDARD: PRT; 1452 AA.
XX
XX xxxxxx
XX
XX
XX
XX
XX Sequence 8, Application US/08087244A
XX
XX Sequence 8, Application US/08087244A
XX Patent No. 5863755
XX GENERAL INFORMATION:
XX APPLICANT: Schlessinger, Joseph
XX APPLICANT: Sap, Jan M.
XX APPLICANT: Ullrich, Axel
XX APPLICANT: Vogel, Wolfgang
XX APPLICANT: Fuchs, Miriam
XX TITLE OF INVENTION: NOVEL RECEPTOR-TYPE PHOSPHOTYROSINE
XX TITLE OF INVENTION: PHOSPHATASE-KAPPA
XX NUMBER OF SEQUENCES: 11
XX CORRESPONDENCE ADDRESS:
XX ADDRESSEE: PENNIE & EDMONDS
XX STREET: 1155 Avenue of the Americas
XX CITY: New York
XX STATE: New York
XX COUNTRY: U.S.A.
XX ZIP: 10036
XX
XX COMPUTER READABLE FORM:
XX MEDIUM TYPE: floppy disk
XX COMPUTER: IBM PC compatible
XX OPERATING SYSTEM: PC-DOS/MS-DOS
XX SOFTWARE: PatentIn Release #1.0, Version #1.25
XX CURRENT APPLICATION DATA:
XX APPLICATION NUMBER: US/08/087,244A
XX FILING DATE: 01-JUL-1993
XX CLASSIFICATION: 435
XX ATTORNEY/AGENT INFORMATION:
XX NAME: Coruzzi, Laura A.
XX REGISTRATION NUMBER: 30,742
XX REFERENCE/DOCKET NUMBER: 7683-042
XX TELECOMMUNICATION INFORMATION:
XX TELEPHONE: 212-790-9090
XX TELEFAX: 212-869-8664/9741
XX TELEX: 66141 PENNIE
XX INFORMATION FOR SEQ ID NO: 8:
XX SEQUENCE CHARACTERISTICS:
XX LENGTH: 1452 amino acids
XX TYPE: amino acid
XX STRANDEDNESS: single
XX TOPOLOGY: unknown
XX MOLECULE TYPE: protein
XX SEQUENCE 1452 AA; 163632 MW; 11201300 CN;
SO

```

```

Db      887  OHITOMKACGEYGEKEEYSEFFEQSAPWDSAKDENRKNRRYGNIIAYDHSRYRLQIE 946
Oy      224  OHLCQLRGALGLALRG-YECFLFEFLGWSPLRRROAVFLEDMESFRTQNSHDLPEFP 282
Db      947  GDTN-SDIINGNYI-DGYHRPNHIIAOG-PMOTIIDFMRMYHENTA 992
Oy      283  GHGEPEBESPCIMESFLRTGATGATESLTKESLLQWWSLSDAPEPS 331

RESULT  12
ID      US-08-449-644-8          STANDARD:          PRT:          1452 AA.
XX
XX      xxxxxx
XX
XX
DE      Sequence 8, Application US/08449644
CC
CC      Sequence 8, Application US/08449644
CC      Patent No. 5856162
CC      GENERAL INFORMATION:
CC      APPLICANT: Schliesinger, Joseph
CC      APPLICANT: Sap, Jan M.
CC      APPLICANT: Ulirich, Axel
CC      APPLICANT: Vogel, Wolfgang
CC      APPLICANT: Fuchs, Miriam
CC      TITLE OF INVENTION: NOVEL RECEPTOR-TYPE PHOSPHOTYROSINE
CC      TITLE OF INVENTION: PHOSPHATASE-KAPPA
CC      NUMBER OF SEQUENCES: 11
CC      CORRESPONDENCE ADDRESS:
CC      ADDRESSEE: PENNIE & EDMONDS
CC      STREET: 1155 Avenue of the Americas
CC      CITY: New York
CC      STATE: New York
CC      COUNTRY: U.S.A.
CC      ZIP: 10036
CC
CC      COMPUTER READABLE FORM:
CC      MEDIUM TYPE: Floppy disk
CC      COMPUTER: IBM PC compatible
CC      OPERATING SYSTEM: PC-DOS/MS-DOS
CC      SOFTWARE: PatentIn Release #1.0, Version #1.25
CC      CURRENT APPLICATION DATA:
CC      APPLICATION NUMBER: US/08/449,644
CC      FILING DATE: 24-MAY-1995
CC      CLASSIFICATION: 514
CC      PRIOR APPLICATION DATA:
CC      APPLICATION NUMBER: US 08/087,244
CC      FILING DATE: 01-JUL-1993
CC      ATTORNEY/AGENT INFORMATION:
CC      NAME: Coruzzi, Laura A.
CC      REGISTRATION NUMBER: 30,742
CC      REFERENCE/DOCKET NUMBER: 7683-042
CC      TELECOMMUNICATION INFORMATION:
CC      TELEPHONE: 212-790-9090
CC      TELEFAX: 212-869-8864/9741
CC      TELEX: 66141 PENNIE
CC      INFORMATION FOR SEQ ID NO: 8:
CC      SEQUENCE CHARACTERISTICS:
CC      LENGTH: 1452 amino acids
CC      TYPE: amino acid
CC      STRANDEDNESS: single
CC      TOPOLOGY: unknown
CC      MOLECULE TYPE: protein
CC      SEQUENCE 1452 AA; 163632 MW; 11201300 CN.

3.ary Match 3.1%; Score 98; DB 2; Length 1452;
B-adj Local Similarity 18.3%; Pred. No. 1.94e+01;
Matches 20; Conservative 37; Mismatches 48; Indels 4; Gaps 4;

```

Db 887 QHITOMKACBEGYGFKEYESFEQOSAPWDSAKKDEMRKRGNIATYDHSRYVLOMLE 946
 QY 224 QHLCOLRGALGLARG-YECFLFEFLGWSPRLRRQAVFLDEMSFRTQSSRDLEFPF 282
 Db 947 GDTN-SDYINGNYI-DGYHRPNHYIATOG-PMOETIYDFWRMWHENTA 992
 QY 283 GHGELPGLGLESPICMESFLRTGAYAGTESLRTKESLLQVWLSLWDAEPS 331

RESULT 13
 ID US-08-991-258A-4 STANDARD: PRT: 1452 AA.

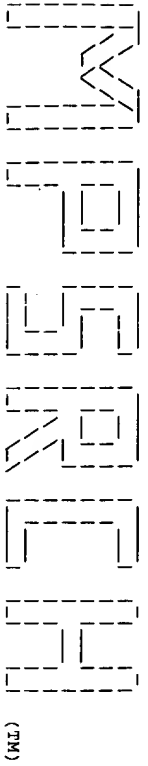
Sequence 4, Application US/08991258A
 Patent No. 5928887
 GENERAL INFORMATION:
 APPLICANT: Cheng, Jili
 APPLICANT: Lasky, Laurence A.
 TITLE OF INVENTION: A NOVEL KAPPA/MU-LIKE PROTEIN TYROSINE
 NUMBER OF SEQUENCES: 10
 CORRESPONDENCE ADDRESSES:
 ADDRESSEE: FLEHR, HOBBACH, TEST, ALBRITTON & HERBERT, LLP
 STREET: 4 Embarcadero Center, Suite 3400
 CITY: San Francisco
 STATE: California
 COUNTRY: United States
 ZIP: 94111
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/991,258A
 FILING DATE: 17-DEC-1997
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/652,971
 FILING DATE: 24-MAY-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Dreger, Walter H.
 REGISTRATION NUMBER: 24,190
 REFERENCE/DOCKET NUMBER: A-63478-3/WHD/MTK
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 781-1989
 TELEFAX: (415) 398-3249
 TELEX:
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1452 amino acids
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 CC SEQUENCE 1452 AA: 163593 MW: 11130193 CN:
 Query Match 3.0%; Score 95; DB 2; Length 1452;
 Best Local Similarity 18.3%; Pred. No. 3.09e+01;
 Matches 20; Conservative 37; Mismatches 48; Indels 4; Gaps 4;
 Db 887 QHITOMKACBEGYGFKEYESFEQOSAPWDSAKKDEMRKRGNIATYDHSRYVLOMLE 946
 QY 224 QHLCOLRGALGLARG-YECFLFEFLGWSPRLRRQAVFLDEMSFRTQSSRDLEFPF 282
 Db 947 GDTN-SDYINGNYI-DGYHRPNHYIATOG-PMOETIYDFWRMWHENTA 992
 QY 283 GHGELPGLGLESPICMESFLRTGAYAGTESLRTKESLLQVWLSLWDAEPS 331

RESULT 14
 ID US-08-652-971-4 STANDARD: PRT: 1452 AA.

Sequence 4, Application US/08652971
 Patent No. 5814507
 GENERAL INFORMATION:
 APPLICANT: Cheng, Jili
 APPLICANT: Lasky, Laurence A.
 TITLE OF INVENTION: A NOVEL KAPPA/MU-LIKE PROTEIN TYROSINE
 NUMBER OF SEQUENCES: 10
 CORRESPONDENCE ADDRESSES:
 ADDRESSEE: Genentech, Inc.
 STREET: 460 Point San Bruno Blvd.
 CITY: South San Francisco
 STATE: California
 COUNTRY: United States
 ZIP: 94080
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/652,971
 FILING DATE:
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Dreger, Ginger R.
 REGISTRATION NUMBER: 33,055
 REFERENCE/DOCKET NUMBER: P1033
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 225-3216
 TELEFAX: (415) 952-9881
 TELEX: 910 371-7168
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1452 amino acids
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 CC SEQUENCE 1452 AA: 163593 MW: 11130193 CN:
 Query Match 3.0%; Score 95; DB 2; Length 1452;
 Best Local Similarity 18.3%; Pred. No. 3.09e+01;
 Matches 20; Conservative 37; Mismatches 48; Indels 4; Gaps 4;
 Db 887 QHITOMKACBEGYGFKEYESFEQOSAPWDSAKKDEMRKRGNIATYDHSRYVLOMLE 946
 QY 224 QHLCOLRGALGLARG-YECFLFEFLGWSPRLRRQAVFLDEMSFRTQSSRDLEFPF 282
 Db 947 GDTN-SDYINGNYI-DGYHRPNHYIATOG-PMOETIYDFWRMWHENTA 992
 QY 283 GHGELPGLGLESPICMESFLRTGAYAGTESLRTKESLLQVWLSLWDAEPS 331
 RESULT 15
 ID US-08-769-399-4 STANDARD: PRT: 1452 AA.
 AC xxxxxx
 CC xxxxxx
 DE Sequence 4, Application US/08759399

XX Sequence 4, Application US/08769399
CC Patent No. 5976852
CC GENERAL INFORMATION:
CC APPLICANT: Cheng, Jill
CC APPLICANT: Lasky, Laurence A.
CC TITLE OF INVENTION: A NOVEL KAPPA/MU-LIKE PROTEIN TYROSINE
CC TITLE OF INVENTION: PHOSPHATASE, PTP LAMBDA
CC NUMBER OF SEQUENCES: 10
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Genentech, Inc.
CC STREET: 460 Point San Bruno Blvd.
CC CITY: South San Francisco
CC STATE: California
CC COUNTRY: United States
CC ZIP: 94080
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/769,399
CC FILING DATE:
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Dreger, Ginger R.
CC REGISTRATION NUMBER: 33,055
CC REFERENCE/DOCKET NUMBER: P1033
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (415) 225-3216
CC TELEFAX: (415) 952-9881
CC TELEX: 910 371-7168
CC INFORMATION FOR SEQ ID NO: 4:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1452 amino acids
CC TYPE: amino acid
CC STRANDEDNESS:
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 1452 AA; 163593 MW; 11130193 CN;
SQ
Query Match 3.0%; Score 95; DB 2; Length 1452;
Best Local Similarity 18.38; Pred. No. 3.09e+01;
Matches 20; Conservative 37; Mismatches 48; Indels 4; Gaps 4;
Db 887 QHITOMKCAEGYGFKEEYESFFEGOSAPMDSAKKDEMRMKNRYGNITAYDHSRYRLQMLE 946
Q 224 QHICQLRGALGLALRG-YECFLFEFLGVWSPFLRRQAVFLQEDMESFRTONSSRDLEFPF 282
947 GDNN-SDYINGNYI-DGYHRPNHYIATQG-PMQETIYDFWRVWMENTA 992
QY 283 GHGELPEGLGSPCIMESFLRTGAVAGTESLRTKESLLQVWSLSMDAEP 331

Search completed: Thu Jan 27 17:40:29 2000
Job time : 18 secs.



Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPArch_PP protein - protein database search, using Smith-Waterman algorithm
on: Thu Jan 27 17:39:02 2000; Maspar time 26.18 Seconds
854.643 Million cell updates/sec

Tabular output not generated.

Title: >US-09-397-558-2
Description: (1-410) from US09397558.pep
Perfect Score: 3201
Sequence: 1 MELPPVLAIRSRVLEAAV.....RDCRMDAGTAVWAGILV 410

Scoring table: PAM 150
Gap 11

Searched: 179066 seqs, 54579741 residues
Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: spiremb19
1:sp-archaea 2:sp-bacteria 3:sp-fungi 4:sp-human
5:sp-invertebrate 6:sp-mammal 7:sp-mhc 8:sp-organelle
9:sp-phage 10:sp-plant 11:sp-rodent 12:sp-unclassified
13:sp-vertebrate 14:sp-virus

Statistics: Mean 48.851; Variance 99.437; scale 0.491

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	224	7.0	585	4	075539	1.27e-18
2	147	4.6	787	10	082362	1.08e-06
3	121	3.8	664	2	052820	1.08e-02
4	115	3.6	238	2	083396	6.60e-02
5	115	3.6	367	8	021203	6.60e-02
6	115	3.6	405	2	083354	6.60e-02
7	111	3.5	247	2	086825	2.13e-01
8	113	3.5	890	2	034431	1.19e-01
9	108	3.4	493	10	023213	5.06e-01
10	110	3.4	591	4	000553	2.85e-01
11	110	3.4	551	4	060227	2.85e-01
12	110	3.4	521	4	013036	2.85e-01
13	107	3.3	521	4	026579	6.71e-01
14	106	3.3	786	2	052499	8.90e-01
15	105	3.3	811	3	087145	1.18e+00
16	101	3.2	248	2	066553	3.54e+00
17	102	3.2	260	8	037371	2.70e+00
18	102	3.2	334	1	029818	2.70e+00
19	102	3.2	362	1	017870	2.70e+00
20	102	3.2	504	4	099624	2.70e+00

21	103	3.2	519	2	033968	POTASSIUM CHANNEL	2.05e+00
22	103	3.2	572	2	P71187	TRBL	2.05e+00
23	102	3.2	599	1	042922	HYPOTHEICAL 65.3 KD P	2.70e+00
24	95	3.1	99	1	028150	HYPOTHEICAL 11.6 KD P	6.07e+00
25	95	3.1	109	2	086493	HYPOTHEICAL 10.8 KD P	6.07e+00
26	96	3.1	172	8	021120	CYTOCHROME OXIDASE I	6.07e+00
27	98	3.1	231	3	074955	TRANSLIN ASSOCIATED PR	7.91e+00
28	93	3.1	261	2	086349	HYPOTHEICAL 27.4 KD P	6.07e+00
29	98	3.1	291	2	086873	INTEGRAL MEMBRANE PROT	7.91e+00
30	96	3.1	298	11	088751	CALDENRIN	7.91e+00
31	98	3.1	309	2	086546	POTASSIUM CHANNEL	7.91e+00
32	98	3.1	322	2	067737	DOLICHO-1-PHOSPHATE MAN	7.91e+00
33	100	3.1	338	2	055226	ACYLTRANSFERASE	7.91e+00
34	98	3.1	393	4	099454	HNPC	7.91e+00
35	98	3.1	443	1	026881	SODIUM/DICARBOXYLATE O	4.64e+00
36	100	3.1	443	1	027326	POTASSIUM CHANNEL RELA	6.07e+00
37	99	3.1	825	4	043525	POTASSIUM CHANNEL HOMO	6.07e+00
38	98	3.1	844	11	054999	PAIRED-IG-LIKE RECEPTO	7.91e+00
39	98	3.1	844	11	043796	VOLTAGE GATED POTASSIU	7.91e+00
40	98	3.1	852	11	088943	POTASSIUM CHANNEL	7.91e+00
41	98	3.1	854	4	075580	NEURONAL DELAYED-RECTI	7.91e+00
42	98	3.1	872	4	043526	POTASSIUM CHANNEL	7.91e+00
43	99	3.1	873	11	088944	POTASSIUM CHANNEL	6.07e+00
44	99	3.1	888	3	003092	SIMILAR TO SMO2_YEAST	6.07e+00
45	98	3.1	2657	5	077380	MAL3P6.15 PROTEIN	7.91e+00

ALIGNMENTS

RESULT	1	PRELIMINARY	PRT	585 AA.
IT	075539			
AC	075539			
DT	01-NOV-1998 (TREMBL:REL. 08, CREATED)			
DT	01-NOV-1998 (TREMBL:REL. 08, LAST SEQUENCE UPDATE)			
DT	01-NOV-1998 (TREMBL:REL. 08, LAST ANNOTATION UPDATE)			
DE	HYPOTHEICAL 63.6 KD PROTEIN (FRAGMENT).			
OS	HOMO SAPIENS (HUMAN).			
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; PRIMATES;			
NC	CARTHINI; HOMINIDAE; HOMO.			
RC	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE-BRAIN;			
FX	REDLINE: 97264341.			
FA	YU W., ANDERSSON B., WORLEY K.C., MUZY D.M., DING Y., LIU W.,			
RT	"Large-scale concatenation cDNA sequencing."			
RT	GENOME RES. 7:353-358(1997).			
RT	[3]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE-BRAIN;			
YU W., GIBBS R.A.;				
RL	SUBMITTED (JUN-1998) TO EMBL/GENBANK/DBJ DATA BANKS.			
DE	EMBL: AF070572; G3387942; -			
KW	HYPOTHEICAL PROTEIN.			
FT	NON TER			
SO	SEQUENCE 585 AA; 63579 MW; 8190CFD1 CRC32;			
Query Match:	7.08; Score 224; DB 4; Length 585;			
Best Local Similarity 100.0%; Pred. No. 1.27e-18;				
Matches: 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
DB: 493 MELPPVLAIRSRVLEAAVYTFMFST 520				
QY: 1 MELPPVLAIRSRVLEAAVYTFMFST 29				

[illegible]

```

RA SOEKIKI A., TACCONI E., TAKAGI T., TAKAHASHI H., TAKEMAU K.,
RA TOSUEUCHI M., TAMAKOSHI E., TANAKA T., TERPETA P., TOGONOI A.,
RA TOSATO V., UCHITAMA S., VANDEBOL M., VANNIER F., VASSAROTTI A.,
RA VIARI A., WAMBUT R., WEDLER E., WEDLER H., WEITENEGGER T.,
RA WINTERS P., WIPAT A., YAMAMOTO H., YAMANE K., YASUMOTO K., YATA K.,
RA YOSHIDA K., YOSHIKAWA H.F., ZUMSTEIN E., YOSHIKAWA H., DANCHIN A.,
RT "The complete genome sequence of the gram-positive bacterium Bacillus
RT subtilis".
RL NATURE 390:249-256(1997).
RN [2]
RN RC SEQUENCE FROM N.A.
RC STRAIN-168;
RA KUNST F., OGASAWARA N., YOSHIKAWA H., DANCHIN A.,
RL SUBMITTED (NOV-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [3]
RN RP SEQUENCE FROM N.A.
RC STRAIN-168;
RA FOULGER D., ERRINGTON J.,
RL SUBMITTED (AUG-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
RA EMBL; 299112; E1185157;
RL EMBL; Y13937; E323521;
DR PFAM; PF00122; E1-E2 ATPase; 1.
SQ SEQUENCE 890 AA; 97292 MW; 319B11CD CRC32;

Query Match 3.5%; Score 113; DB 2; Length 890;
Best Local Similarity 29.5%; Pred. No. 1,19e-01;
Matches 31; Conservative 24; Mismatches 42; Indels 8; Gaps 7;

Db 696 LASANGELIVMFAFLALPLPLVPIQILVNWNLDTGPAALAMDPEEGYMKRRKPRHP 755
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
Oy 72 LQPVKKQVLYLLGALLSMALQDLDR-HGLM-NLGLPSL--FALDILATATWRYSVRRRH 127
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
Db 756 KEGVARFLGKWKVSRGFLIGAVATLAEIIVYHRRNPENLAYAQTI 800
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
Oy 128 YPTWRRL-F-YLCPGSLIAGSAVLTAFAFETRD-NF-YIHSI 168

RESULT 9 PRELIMINARY; PRT; 493 AA.
ID 023213;
AC 023213;
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE MYO-INOSITOL TRANSPORT PROTEIN HOMOLOG.
OS ARABIDOPSIS THALIANA (MOUSE-EAR CRESS).
OC EUPHYLLIPHYTES; STREPTOPHYTA; EMBRYOPHYTA; TRACHEOPHYTA;
OC EUPHYLLIPHYTES; SPERMATOPHYTA; MAGNOLIOPHYTA; EUDICOTYLEDONS; ROSIDAE;
CAPRALES; BRASSICACEAE; ARABIDOPSIS.
[1]
SQ SEQUENCE FROM N.A.
RA BEVAN M., TERRYN N., VOS P., HEINEN L., MEMES H.W., SCHUELLER C.,
RA CHALMAZIS N.,
RL SUBMITTED (OCT-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [2]
RN RP SEQUENCE FROM N.A.
RA EU ARABIDOPSIS SEQUENCING PROJECT ESSA;
RL SUBMITTED (SEP-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; 299708; E353127;
DR PFAM; PF00083; sugar_tf; 1.
SQ SEQUENCE 493 AA; 52922 MW; CF55253A CRC32;

Query Match 3.4%; Score 108; DB 10; Length 493;
Best Local Similarity 29.3%; Pred. No. 5.06e-01;
Matches 12; Conservative 16; Mismatches 13; Indels 0; Gaps 0;

Db 62 IINLCALVSGSLAGRTSDIIGRRYTVLASLTFMLGSLMG 102
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
Oy 49 VLOCFDPLGSLSVWTVIAMARLQPVKKQVLYLLGALLS 89

RESULT 10 PRELIMINARY; PRT; 591 AA.
ID 000553;
AC 000553;

```

DT	01-JUL-1997	(TREMBLREL. 04, CREATED)
DT	01-JUL-1997	(TREMBLREL. 04, LAST SEQUENCE UPDATE)
DT	01-NOV-1998	(TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE	FOLATE CARRIER.	
OS	HOMO SAPIENS (HUMAN).	
OC	EUDAROTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; PRIMATES;	
CC	CATARRHINI; HOMINIDAE; HOMO.	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	TISSUE=SMALL INTESTINE;	
RA	MEDLINE; 97193655.	
RA	NGUYEN T.T., DYER D.L., DUNNING D.D., RUBIN S.A., GRANT K.E.,	
RA	SAID H.M.:	
RT	"Human intestinal folate transport: cloning, expression, and	
RL	distribution of complementary RNA."	
DR	GASTROENTEROLOGY 112:783-791(1997).	
SO	EMBL; AF004354; G2209135; ..	
	SEQUENCE 591 AA; 64868 MW; 39146F8C CRC32;	
 Query Match 3.4% ; Score 110; DB 4; Length 591;		
Best Local Similarity 31.8%; Pred. No.2.85e-01;		
Matches	28; Conservative	22; Mismatches 29; Indels 9; Gaps 9;
 Db 192 FLTFSYVALFLKPKRSLEP-NDDRGRCETASLEFRANPGGGLGHALRYACG-DS 249		
Qy	243 FLEFFLGWSPLLR-RQAVFLEDWESPRONSSRDLEPF-PGHGE-LPGLESPTMES 299	
 Db 250 VLARMRELG-DSLR-R-POLRLMSLWM 274		
Qy	300 FL-RTGAAGTESLRKESLLQVWSLSM 326	
 RESULT 11 PRELIMINARY; PRT; 591 AA.		
ID	060227;	
AC	060227;	
DT	01-AUG-1998 (TREMBLREL. 07, CREATED)	
DT	01-AUG-1998 (TREMBLREL. 07, LAST SEQUENCE UPDATE)	
DT	01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)	
DE	REDUCED FOLATE CARRIER.	
GN	RFCL.	
OS	HOMO SAPIENS (HUMAN).	
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; PRIMATES;	
CC	CATARRHINI; HOMINIDAE; HOMO.	
RM	[1]	
RP	SEQUENCE FROM N.A.	
RC	TISSUE=LIVER;	
RA	TOLNER B., ROY K., SIROTNAK F.M.;	
RI	SUBMITTED (MAR-1997) TO EMBL/GENBANK/DBJ DATA BANKS.	
DR	EMBL; U92873; G2967654; ..	
DR	EMBL; U92869; G2967654; JOINED.	
DR	EMBL; U92870; G2967654; JOINED.	
DR	EMBL; U92871; G2967654; JOINED.	
DR	EMBL; U92872; G2967654; JOINED.	
SO	SEQUENCE 591 AA; 64871 MW; 5C105C56 CRC32;	
 Query March 3.4% ; Score 110; DB 4; Length 591;		
Best Local Similarity 31.8%; Pred. No.2.85e-01;		
Matches	28; Conservative	22; Mismatches 29; Indels 9; Gaps 9;
 Db 192 FLTFSYVALFLKPKRSLEP-NDDRGRCETASLEFRANPGGGLGHALRYACG-DS 249		
Qy	243 FLEFFLGWSPLLR-RQAVFLEDWESPRONSSRDLEPF-PGHGE-LPGLESPTMES 299	
 Db 250 VLARMRELG-DSLR-R-POLRLMSLWM 274		
Qy	300 FL-RTGAAGTESLRKESLLQVWSLSM 326	
 RESULT 12 PRELIMINARY; PRT; 591 AA.		
ID	Q13026	
AC	Q13026;	
DT	01-NOV-1996 (TREMBLREL. 01, CREATED)	
DT	01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)	

DT	01-NOV-1998 (TREMABEL. 08. LAST ANNOTATION UPDATE)
DE	65 KDA HYDROPHOBIC PROTEIN.
OS	HOMO SAPIENS (HUMAN).
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUOTHERIA; PRIMATES;
OC	CATARRHINI; HOMINIDAE; HOMO.
RN	[1]
RP	SEQUENCE FROM N.A.
RC	TISSUE-LYMPHOMA;
RC	MEDLINE: 95155380.
RX	WILLIAMS F.M., FLINTOFF W.F.;
RT	"Isolation of a human cDNA that complements a mutant hamster cell
RT	defective in methotrexate uptake."
RL	J. BIOL. CHEM. 270:2987-2992(1995).
SD	EMBL: U17566; G965000; -
DR	SEQUENCE 591 AA; 64852 MW; 930D15BA CRC32;

```

Query Match:          3.4%; Score 110; DB 4; Length 591;
Best Local Similarity 31.8%; Pred. No. 2,85e-01;
Matches 28; Conservative 22; Mismatches 29; Indels 9; Gaps 9;

Db 192 FLTSPVLALEFLKPKRSITFF-NNDDRGRGRETSAEIERNNPPGGKLGHALRVAGG-DS 249
|||:::||||:::||||:::|||||||:::|||||
OY 243 FLEFELCVMSPLRR-RQAVFELEDWESSRYONSSRDLEPF-PGHGE-LPEGLSEPCIMES 299
||||:::||||:::||||:::|||||||:::|||||
Db 250 VIARMLRELG-DLSLR-R-POLRLSLWM 274
| |||:::|||||
OY 300 FL-RTGAIVAGTESLRTKESLIOWWSLSW 326
```

RESULT	13			
ID	Q26579	PRELIMINARY	PRT	521 AA.
AC	Q26579;			
DT	01-NOV-1996 (TREMBL,REL. 01, CREATED)			
DT	01-NOV-1996 (TREMBL,REL. 01, LAST SEQUENCE UPDATE)			
DT	01-NOV-1998 (TREMBL,REL. 08, LAST ANNOTATION UPDATE)			
DE	GLUCOSE TRANSPORT PROTEIN.			
GN	GTP1.			
OS	SCHISTOSOMA MANSONI (BLOOD FLUKE).			
OC	EUKARYOTA; METAZOA; PLATHELMINTHES; TREMATODA; DIGENEA; STRIGEIDIDA;			
OC	SCHISTOSOMATOIDEA; SCHISTOSOMATIDAE; SCHISTOSOMA.			

RP SEQUENCE FROM N.A.
RC STRAIN-PUERTO RICAN;
RX MEDLINE; 9410848.
RA SHELLEY P. J., KIM J.W., CUNNINGHAM J.M., SHOEMAKER C.B.;
RT "Cloning, characterization, and functional expression of cDNAs
encoding glucose transporter proteins from the human parasite
Schistosoma mansoni.";
RD J. Biol. Chem. 269:4247-4253(1994).
DR EMBL; L25065; G407041; -.
SO PR00083; sugar_tr; 1.
SEQUENCE 521 AA; 56815 MW; 3C17EEEB CRC32;

```

Query Match      3.3%  Score 107;  DB 5:  Length 521;
Best Local Similarity 31.3%  Pred. No. 6,718-01;
Matches 15;  Conservative 18;  Mismatches 14;  Indels 1;  Gaps 1;

Db 303 DYTATGTCVCAIGVNLNVTYVVSLEPLIRACGRITLLMPTVSLALSTLL 350
      |||.....:.....:.....:.....:.....:.....:.....:.....:
48 DV-LQFCDFGLSLMSVNTVTIAAMRLQPVKKQVYLYLLGAILLSMALQI 94

```

RESULT	14		
ID	05249	PRELIMINARY:	PRT; 786 AA.
AC	05249:		
DT	01-JUN-1998	(TREMBLER, 06,	CREATED)
DT	01-JUN-1998	(TREMBLER, 06,	LAST SEQUENCE UPDATE)
DT	01-AUG-1998	(TREMBLER, 07,	LAST ANNOTATION UPDATE)
DE	INTEGRAL MEMBRANE PROTEIN.		
GN	EXP.		
OS	BRADYRHIZOBIIUM JAPONICUM.		
OC	BACTERIA: PROTEOBACTERIA: ALPHA SUBDIVISION: BRADYRHIZOBIIUM GROUP:		
CC	BRADYRHIZOBIIUM.		

F. F. [1]
 P. P. SEQUENCE FROM N.A.
 R. R. STRAIN-35DA 110SPC;
 R. R. BECKER R.U. KOSCH K.
 R. R. * SUBMITTED (DEC 1992) PARINISKE M., MULLER P.,
 DR. EMBL, AF0393906; G2921558; -
 SEQUENCE 786 AA; 84633 MW; 44935912F CRC32;

	Query Match	Similarity	40.0%	Score 106	DB 2	length 786
	Best Local	Similarity	40.0%	Pred. No. 8.90e-01		
	Patches	Conservative	24	Mismatches 17	Indels 7	Gaps 6
D6	541	541	541	541	541	541
07	69	69	69	69	69	69

RESULT	15	PRELIMINARY:	PRT:	811 AA.
ID	P87145			
AC	P87145			
DT	01-AUG-1998 (TEMBELREL, 07, CREATED)			
DT	01-AUG-1998 (TEMBELREL, 07, LAST SEQUENCE UPDATE)			
DT	01-AUG-1998 (TEMBELREL, 07, LAST ANNOTATION UPDATE)			
DE	HYPOTHEETICAL 92.5 KD PROTEIN C25H.03 IN CHROMOSOM			
G4	SPBC25H2.03.			
G5	SCHIZOSACCHAROMYCES POMBE (FISSION YEAST).			
M2	EURK107A: FUNGI; ASCOMYCOTA; ARCHIZASCOMYCETES;			
CC	SCHIZOSACCHAROMYCETALES; SCHIZOSACCHAROMYCETACEAE.			
KN	SCHIZOSACCHAROMYCES.			
	[1]			

RA. DIRSO G. IVE G., BOWMAN S., CHURCH C., WOOD V., BARRELL B.G.,
RA. RAJANIREM M.A., CONNOR R.E.;
RL. SUBMITTED (MAY-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
CC. -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL)
CC. -1- SIMILARITY: TO YEAST YLR386W.
DR. EMBL: 295397; E316120; -
KW. HYPOTHETICAL PROTEIN; TRANSMEMBRANE.
FT. TRANSMEM 58 78 " POTENTIAL.
FT. TRANSMEM 234 254 POTENTIAL.
FT. TRANSMEM 362 382 POTENTIAL.
FT. TRANSMEM 602 622 POTENTIAL.
FT. TRANSMEM 630 650 POTENTIAL.
SQ. SEQUENCE #11 AA: 92452 MW: 112377 CRC32;

Query Match	3.3%;	Score 105;	DB 3;	Length 811;
Be% Local Similarity	23.7%;	Pred. No. 1.18e+00;		
Mismatches	31;	Mismatches	52;	Indels 9;
				Gaps 9;

```

Dc 579 LETPAELDELEKRR-LKOSAP-KLONITFTTLYTAC-NSNIAVRS-L-CLISONEVHANL 6333
G 1 ME:PFVVALISRSVLEAAVYFTFMFSTFHA--DDPGIIVECIMDYDLOCDFEGL 59
Db 634 LSVAFHIFENIDMLIQDKNLYQLIESVFTYMRQLLEPEKYPILHAKLYGILMLPOSS 6933
Cy 60 MSVAVY-Y-IAARAPVVKOVLYLLGLMLSLMALQDRHGILMLLGPSTLA-LEIATAW 117
Dd 634 AFRITZRLQOC 704
Dy 118 TVRQVRRR-HC 127

```

Search completed: Thu Jan 27 17:39:53 2000
Job time : 51 secs.

Page Blank (uspto)

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ **BLACK BORDERS**
- ☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☐ **FADED TEXT OR DRAWING**
- ☐ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☐ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☐ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER:** _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.

BEST AVAILABLE COPY